

DOCUMENT RESUME

ED 160 219

PS 010 135

TITLE Current Issues in the Development Project for Head Start Profiles of Program Effects on Children.

INSTITUTION Mediax Associates, Inc., Westport, Conn.

SPONS AGENCY Administration for Children, Youth, and Families (DHEW), Washington, D.C.

PUB DATE 1 Aug 78

CONTRACT REW-105-77-1006

NOTE 51p.

EDRS PRICE MF-\$0.83 HC-\$3.50 Plus Postage.

DESCRIPTORS *Child Development; Data Analysis; *Early Childhood Education; *Evaluation Criteria; Handicapped Children; *Measurement; Measurement Goals; Measurement Techniques; *Program Evaluation

IDENTIFIERS *Project Head Start

ABSTRACT

This collection of 3 papers summarizes the current thinking of those involved in the Head Start Program Effects Measurement Project regarding selected issues in the assessment of Head Start program effects on children's development. The first paper describes the project panelists' current approaches to the selection of measurement dimensions. Criteria for the selection of dimensions are listed, broad constructs or domains of behavior to be assessed are identified, issues relating to the collection of additional data on the various contexts affecting the child are discussed, and technical problems in the measurement of program effects are summarized. The second paper discusses implications for the measurement of program effects on special needs children. Specific topics examined include: (1) legal provisions for including handicapped children in Head Start, (2) definition of the term handicapped, (3) current enrollment of handicapped children in Head Start, (4) concern over labeling handicapped children, (5) Head Start classification system for handicapped children, and (6) options for handling the handicapped in the program effects measurement program. The third paper discusses two approaches to the aggregation of data which could yield suitable indices for the various constructs to be assessed. The judgmental and the empirical approach to weighting information are described and the empirical approach is recommended for the present study. (JMB)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.



HEAD START

Current Issues
in the
Development Project for
Head Start Profiles of Program Effects on Children

August 1, 1978



"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

ACYF

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

Prepared by:

MEDLAX ASSOCIATES, INC.
WESTPORT, CONN. 06880

For the

ADMINISTRATION FOR CHILDREN, YOUTH AND FAMILIES
CONTRACT NO. HEW-105-77-1006

Current Issues
in the
Development Project for
Head Start Profiles of Program Effects on Children

HEW-105-77-1006

August 1, 1978

PS 010135

prepared by
Mediav Associates
21 Charles Street
Westport, Conn. 06880

Table of Contents

I. Current Approaches of Panelists	1
A. Purpose of Paper	2
B. Criteria for the Selection of Dimensions	3
C. Initial Approaches to the Selection of Constructs	7
a. Status Constructs: Basic Resources	
1. Physical Development	9
2. Social Development	10
3. Cognitive Development	11
Notes on Status Constructs: Basic Resources	11
b. Status Constructs: Attitudes and Feelings	
4. Emotional Development	12
5. Aesthetic Moral/Ethical Development	12
Notes on Status Constructs: Attitudes and Feelings	12
c. Application Constructs	
Introduction to Application Constructs	13
6. Communications Competence	13
7. Specific Functional Competence	14
Note on Graphic Presentation of Proposed Constructs	14
Graphic - Tentative Proposed Constructs	15
D. Implications for the Evaluation of Program Effects	16
1. The Environment of the Data-Gathering Situation	16
2. The Measurement of Competencies in a Variety of Contexts	17
3. The Environment of the Program Being Evaluated	17
4. The Family Environment	17
5. Community Environment	18
E. Technical Problems	20
1. The Measurement of Affective Variables	20
2. Techniques for Combining Affective and Cognitive Dimensions	20
3. New Approaches and Techniques to Measure Stress	20
4. Profiling Combinations of Constructs	20
5. New Technologies	21
6. Allowing for Future Research Findings	21
7. Assuring Consideration of Local Priorities and Standards	21
8. Allowing for a Wide Variety of Program Practices and Theories	21
9. Ethical Considerations	22
F. Current Assignments of Panelists	23
II. A Discussion of Implications for Measurement of Program Effects on Special Needs Children	27
III. Approaches to the Aggregation of Information	37

I. Current Approaches of Panelists

A. PURPOSE

The purpose of these papers is to summarize current thinking of those involved in the Head Start Program Effects Measurement Project. They represent individual ideas obtained from consultants to the project, project panelists, ACYF staff and the project staff. Where possible, sources have been identified; but in many cases, origins have been lost over time as ideas have become integrated into day-to-day thinking of the project. This summary is intended for review and discussion so as to facilitate a coalescing of thinking among all those involved in the project. It is an expression of the original commitment of the project to document proceedings so that all those affected by the project's outcomes can understand and participate in their development.

The many scholars and others involved in this project have devoted much of their lives to the ideas contributing to this project and to this brief statement of preliminary thinking. To them must go the credit for what is of value. The need for a summary, however, precludes exhaustive research and documentation of the statements herein. There are probable errors and there are many failures to give credit to sources. It would be unfair to attribute such defects to anyone other than the writer. We urge those reading this document to advise us on how it and the project's process may be improved.

Herman Taub, Project Director
July 30. 1978

B. CRITERIA FOR THE SELECTION OF DIMENSIONS

A great deal of the project's efforts to date have been devoted to defining the dimensions of child development. The literature has been reviewed, scholars' opinions have been surveyed and, most importantly, an extensive and detailed compilation of desired attainments has been obtained from some 378 parents, practitioners, and public school teachers representing the diverse populations served by Head Start. The results of these investigations have been reported separately.

An initial problem confronting the project's panelists is to select from among these comprehensive compilations of dimensions of child development, those which are to be focused on in the outcomes of the project (and for which measuring techniques are to be subsequently provided). Some criteria for this selection are developing along the following lines:

1. There must not be so few dimensions as to mask independent factors and significantly different factors in individual profiles of development or too many to facilitate comprehension of their overall significance. By these criteria the panelists felt that a single construct comparable to "IQ" is too few, and the twenty nine (29) recommended by Anderson & Messick are too many.
2. The constructs selected must have face validity as being important to child development. They must be understood and accepted by Head Start parents, practitioners, and policy makers as well as by the

community of scholars and researchers. They must have "consumer acceptance" with all these constituencies.

3. The dimensional constructs must be useful for their intended purposes:
 - a. to inform policy makers
 - b. to inform Head Start administrators and communities of "reasonable expectations."
 - c. to determine the relative effectiveness of alternative intervention strategies.
 - d. to provide data on long-range program effectiveness.
4. Although the outcomes of this project are intended to measure program effects, rather than to evaluate individual children, they will inevitably result in focusing on similar dimensions in individual children. The selected dimensions, therefore, must also be appropriate for measures of individual child development, for describing diverse values of child development, and for prescribing the objectives for diverse program approaches.

NOTE: The measuring techniques adopted out of this project will not be directly usable with individual children. Chance errors that can greatly affect measurement of individual children are cancelled out in measuring large numbers of children. Therefore, techniques such as would be developed on this project for program effects would be unreliable for individual children due to the anticipated variability of results. Similarly, program effects would not be valid with individual children, also because validity is dependent on score variability.

5. The selected dimensional constructs must include all factors considered critical to child development. They must be constructed so as to give appropriate weight to their relative importance. The dimensional constructs must be sensitive to all the intended effects of Head Start programs on children. They must include all significant contributors to children's "everyday effectiveness in dealing with both present environment and later responsibilities in school and life." The panelists recognize that there is no rigorous approach to selecting or constructing dimensions and that the objective, and the need, is for practical, useful measures that can command wide respect and application.
6. The constructs selected should reflect both the personal development (what he/she can do) of children and the appropriateness of the behavior in specific situations encountered in the environment of their culture (what he/she does).
7. In as much as we are inevitably involved in evaluations of behavior in defining and measuring constructs, we must consider the interrelationship of affective, cognitive and physical attributes. A child is an inseparable unity, and all behavior involves interactions of factors in all these domains; therefore, the evaluation of program

effects, and the dimensional constructs selected for measurement, must include all these domains.

Although it is popular to describe all classification schemes as "arbitrary abstractions," the panelists feel that application of these criteria will define a set of dimensions which is particularly appropriate to measure the effects of intervention programs on children between the ages of 3 and 7.

C. INITIAL APPROACHES TO THE SELECTION OF CONSTRUCTS

The panelists have indicated some initial approaches to the selection of behavioral constructs (or "domains"). These approaches are summarized here to facilitate communications among panelists and others interested in the project. The process of communication is undoubtedly imperfect and these descriptions no doubt are incomplete in some instances and in error in others; nevertheless, they are presented without further refinement in order to initiate needed comments, discussion, and suggestions.

In accordance with the criteria for selection, several broad constructs (or "domains") of behavior are here conceived--along with their respective dimensions--as the basis for measuring program effects. Some of them reflect the status of children, including such basic resources as physical, cognitive, and social development, together with attitudes and feelings involving these basic resources. Others reflect the application of these resources as competencies in specific contexts, such as appropriate communicative behaviors and appropriate functioning in specific situations (where, incidentally, all measurements must proceed).

The panelists recognize that these constructs are not independent or mutually exclusive; they clearly overlap. The dimensions initially proposed as constituting a construct need to be further defined and verified empirically during the measurement-development phase of this project (Phase II). For the initial conception of constructs, and possibly in the final recommendations, a dimension may be included as a factor in more than one construct.

The panelists are very concerned about misinterpretations of measurement and have defined specific concerns about the use

of the project's outcomes in evaluations (see section IV) which need to be considered.

With these approaches in mind, the following constructs have been initially defined. They will require further consolidation and clarification by the panelists to meet their criteria for selection (see section II).

The constructs are organized as follows:

a. Status Constructs: Basic Resources

1. Physical development
2. Social Development
3. Cognitive Development

b. Status Constructs Attitudes and Feelings

4. Emotional development
5. Aesthetic - moral/ethical development

c. Application Constructs

6. Communication Competence
7. Specific Functioning Competence

a. Status Constructs: Basic Resources

1. Physical Development

(See attainments identified by Head Start parents and teaching staffs for input workshop dimensions 1 thru 5 for specifics defining this construct.)

This construct includes medical health, dental health, nutrition, gross and fine motor development, the availability of ongoing health services and the ability to function under stress (when stress becomes distress). The medical advisors to the panel emphasized that stress is a major cause of physical and psychological disfunction and are particularly significant factors in families living under conditions of high stress.

The affective component of this construct relates to (a) anticipation of reinforcement (positive-negative, intrinsic-extrinsic) and (b) perception of probable success.

2. Social Development

(See attainments identified by Head Start parents and teaching staff for input workshop dimensions 6, 11, 12, 13, 15, 16 and 17, for specifics defining this construct).

This construct includes understanding of feelings of others, ability to get along with peers and others, ability to share and compete effectively, knowledge and skills of own culture and of other important local cultures.

The affective component of this construct includes (a) anticipation of reinforcement (positive-negative, intrinsic-extrinsic) (b) perception of probable success, and (c) attitudes to peers and adults, (also part of emotional development and moral-ethnical development).

3. Cognitive Development

(See attainments identified by Head Start parents and teaching staffs for input workshop dimensions 20 thru 28 for specifics defining this construct.)

This construct includes basic cognitive functions (e.g., memory, perceptual skills, spatial reasoning), learning styles (e.g., field dependent-independent, impulsivity-reflectivity, attention), content (e.g., language, number concepts, knowledge of the worlds of nature, technology, commerce and societal organization), and strategies for coping with low and high rates of change of sensory inputs.

The affective component of this construct includes (a) anticipation of reinforcement (positive-negative, intrinsic-extrinsic), and (b) perception of probable success.

Note On Status Constructs: Basic Resources

Although panelists (and parents, Head Start staffs, etc.) are concerned about undue pressure on children or programs to develop maximally in all dimensions, the above constructs are characterized by a single uni-directional dimension of development (basically "more is better"), e.g., more health, more gross and fine motor skills, greater stress tolerance, more understanding of feelings of others, greater abilities in getting along with others, more styles of learning, more cognitive knowledge and skills, etc...

b. Status Constructs: Attitudes and Feelings

4. Emotional Development

(See attainments identified by Head Start parents and teaching staffs for input workshop dimensions 7 thru 10 for specifics defining this construct.)

This construct includes development of basic feelings of security, love, ability to express feelings, concepts of self and others, of independence, assertiveness and responsibility. These are largely affective dimensions, although they are greatly influenced by cognitive factors (knowledge and skills).

5. Aesthetic - Moral/Ethical Development

(See attainments identified by Head Start parents and teaching staffs for input workshop dimensions 29 thru 32 for specifically defining this construct.)

This construct includes concepts of quality, beauty, standards of relationships with others and values of personal conduct and of the conduct of others.

Notes on Status Constructs: Attitudes and Feelings

Dimensions four (4) and five (5) cannot be evaluated in absolute terms (e.g., more is not necessarily better); they involve a very large number of healthy, "balanced" conditions greatly influenced by local culture and physical environment.

Introduction to Application Constructs

The following constructs are largely dependent upon the above five constructs. They are characterized by their measurement being a function of "appropriateness" to context and perceived purpose. They represent the child's ability to apply selected competencies for a specific purpose and in a specific context. Although a large number of abilities for application could be identified for measurement, the panelists are tentatively identifying the following situational competencies as being particularly significant for 3 to 7 year old children, (see 6 and 7 below).

C. Application Constructs

6. Communications Competence

This construct involves the appropriate application of verbal and non-verbal communication skills in specific contexts (e.g., with peers, family, teacher, etc.) towards a specific purpose. It involves children's abilities both to receive and understand messages of others and to express their own feelings, needs and ideas. A specific skill involved in this competence is code-switching. The affective components involved in this construct combine in a factor of motivation, which is of particular importance to children.

7. Specific Functional Competence

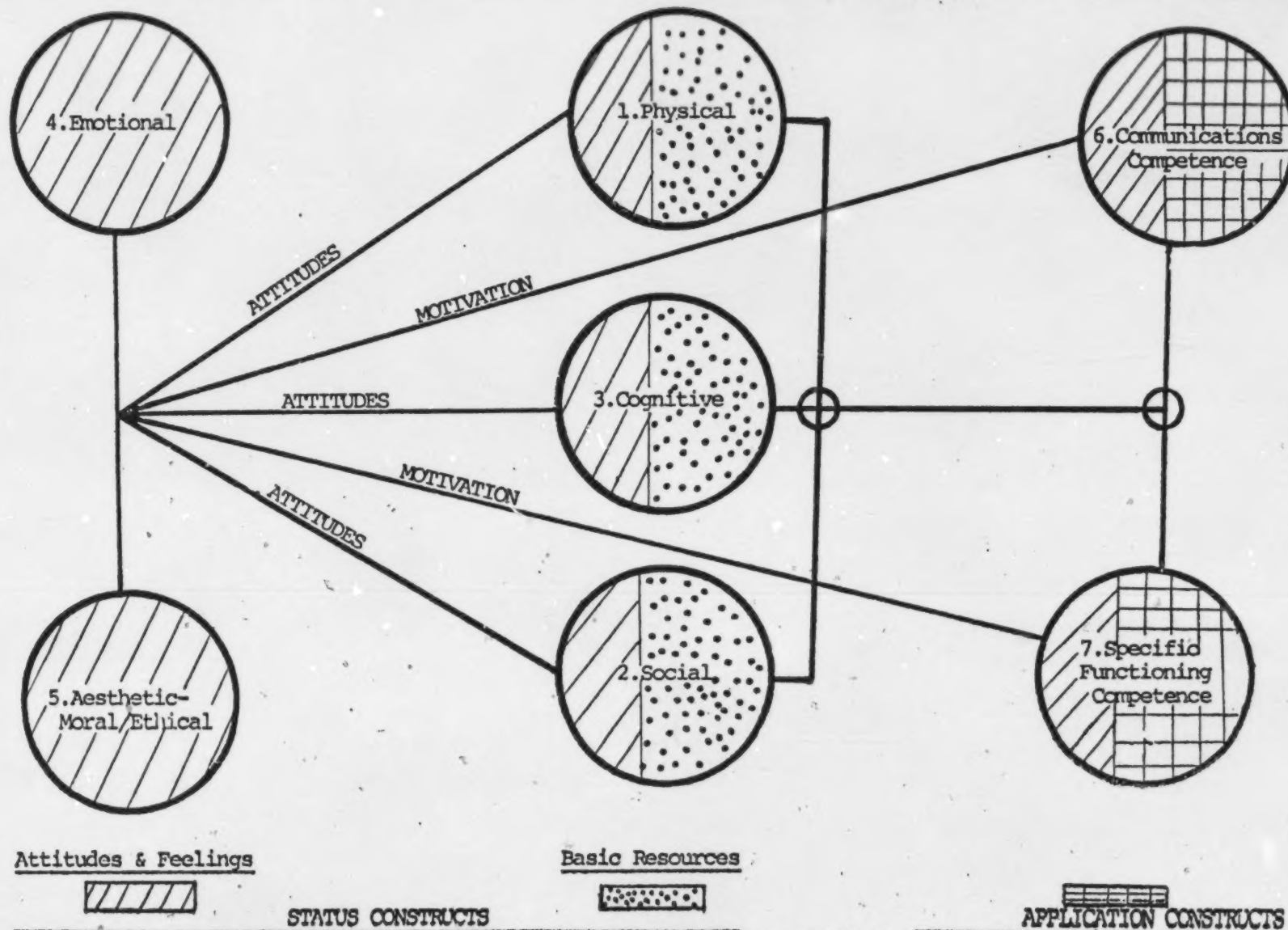
(See attainments identified by Head Start parents and teaching staffs for input workshop dimensions 11 thru 19 for specifically defining this construct.)

This construct includes the application of appropriate knowledge, skills, and attitudes--in games, in playing with peers and adults, in every-day functions in the community, and in Head Start and other school situations. This construct involves the appropriateness of children's behavior in each of these specific contexts to a specific purpose and involves special skills, such as identifying alternative courses of action and the evaluation of their consequences.

The affective components involved in this construct also combine in a dimension of motivation in these specific contexts and toward these specific purposes.

Note on Graphic Representation of Proposed Constructs

On the following page a graphic representation of the functional interrelationships of these seven constructs has been attempted. The representation is proposed to facilitate understanding of the interrelationships of the constructs there tentatively proposed for measurement; the presentation is not intended to represent functioning within a child.



STATUS CONSTRUCTS

Basic Resources

APPLICATION CONSTRUCTS

Figure 1 - Tentative Proposed Constructs of Child Development
for
Head Start Program Effects Measurement Project

D. IMPLICATIONS FOR THE EVALUATION OF PROGRAM EFFECTS

The panelists feel that in using constructs of child development such as proposed herein, a comprehensive profile of program effects on children is possible. For proper evaluation, however, the information-gathering design must also include data on:

- a. The environment of the data-gathering situation.
- b. The measurement of competencies in a variety of contexts.
- c. The environment of the program being evaluated.
- d. The family environment.
- e. The community environment.

The panelists feel that each of these recommendations is critical to a proper evaluation of program effects.

1. The Environment of the Data-Gathering Situation

Behavior in the data-gathering situation, as in all situations, is a function of affective and cognitive responses to the environment. Information on the environment of data-gathering situations is essential for an interpretation of observed behavior. Different environments should be expected to elicit different behaviors.

The panelists therefore feel that it is most important to "test" children in environments that are natural to them for the observed behavior, otherwise, unpredictable adjustment behaviors might be elicited. Towards this objective the panel is investigating the use of standard "play environments" for data gathering across many dimensions.

2. The Measurement of Competencies in a Variety of Contexts

Communications and functioning competencies, such as identified by the panelists, embody affective and cognitive factors appropriate to a particular context. Children will and should behave very differently at home, in the community, and at school. Appropriate behavior in one context cannot be assumed to predict appropriate behavior in another context. Moreover, the purpose of the child in each of these contexts must be considered in evaluating observed behaviors.

3. The Environment of the Program Being Evaluated

Names of programs should not be confused with types of programs. It is useless to evaluate the effects of Head Start, Kindergarten, or Primary Grade programs, or the effects of one in combination with those of another, unless the critical variables describing each program's environment are carefully documented. The panel feels that a significant portion of the program effects variable is contributed by these factors. Towards this need, the panel plans to recommend to ACYF a study and development of measures of program-environment variables.

Preliminary inclinations of panelists include documentation of information on teacher style, competency, and other professional activities; on class size; and on content and style of program activities.

4. The Family Environment

All of the panelists feel that it is most important to obtain data on the family environment in evaluating program effects.

Child development is largely determined by family influences and can only be modified or reinforced by program effects. An evaluation of program effects on children must therefore be considered as a function of the family environment.

Recognizing the importance of the family, ACYF has targeted major objectives of Head Start programs to outcomes affecting the family. Information on the family environment and program effects on the family environment need to be considered in evaluating program effects on children.

The panelists recognize, as does ACYF, that family variables must be considered in evaluating the overall effects of programs and of specific programs on children.

5. Community Environment

The community environment as well as the family environment affects the behavior of children. A child from a minority culture in a strange community environment is impacted with very different needs, influences and effects than a child of the majority culture in a community. (The same is true of a child of a majority culture injected into a minority environment.) Also differences in community economics, practices and social-service resources influence differences in child development.

ACYF has recognized the importance of community influences on child development. The panelists, and ACYF, recognize that information of the community environment should be considered in evaluating the overall effects of programs and specific program effects on children.

The panelists are confident that their approach to the measurement of program effects on children, including data on the contexts affecting the child, will result in a really significant improvement in the evaluation of early child development programs and, as a consequence, the increased effectiveness of early child development programs.

E. TECHNICAL PROBLEMS

The panelists recognize that there are many technical problems involved in attaining the project's outcomes. They, and MediAx, are committed to having these problems resolved. The recommendations and involvement of all available expertise and resources will be considered in seeking solutions.

Some of the presently recognized problems to be resolved in this project are here summarized for attention:

1. The measurement of affective variables, such as motivation (see construct 4), attitudes (see constructs 1, 2 and 3), and aesthetic/moral/ethical development (construct 5) are relatively underdeveloped fields. New approaches and techniques need to be considered and developed.
2. Techniques for combining affective and cognitive dimensions of each of the constructs need to be developed.
3. The panelists feel that variables such as "coping with low and high rates of change in sensory impacts" (construct 3) and the related factor "functioning under stress" (construct 1) are important factors in child development. New approaches and techniques need to be considered and/or developed for these factors.
4. An effective approach to profiling combinations of constructs needs to be developed. Weighting and scaling of individual constructs and dimensions of factors within constructs needs to be considered.

5. The state of the measurement art is advancing rapidly. Long-range perspectives must be considered at each point in making each decision. The advantages of new technologies (computers, electronics, sensing, video taping) must be balanced with their risks.
6. The individual constructs and their profiles must allow for future research findings while retaining their validity as longitudinal measures of program effects on child development.
7. Since many of the constructs are context dependent, appropriate contexts and standards of appropriate behaviors must reflect local customs and practices. Children's competencies must be judged in terms of the local contexts in which, in fact, the children's ultimate "success" will be determined. Techniques need to be developed to assure consideration of local priorities and standards and to avoid imposition of a "foreign message" (e.g. from Washington). The panelists feel that programs and program-effects measures must relate to local priorities in order to have desired effects on child development.
8. The project's outcomes must be validated for use with a wide variety of program practices and a wide variety of theories of child development. The outcomes must not be "unfair" or discourage program innovations. The measures must not be unfairly biased for or

against specific theories of child development, or so fragmented so as to miss the essence of what is "the best" in any approach to child development.

9. The panelists feel that program-effects measures, in particular, must be so developed and implemented as to protect children against abusive or negative experiences. (e.g., children must not be exposed to feelings of "failure" or excessive stress in the testing situation). These ethical considerations, requirements for appropriate psychomatic properties (validity, reliability, etc.), and the practical problems of skilled and efficient administration (e.g., administration time, training of local administrators, ease of evaluation of results, etc.) require that the panelists investigate and evaluate a wide range of alternative technical approaches.

F. CURRENT ASSIGNMENTS OF PANELISTS

Saundra Aldridge

Will provide input to panelists on the concerns of the Head Start community and the requirements of Head Start children, and also to communicate the objectives of the project to Head Start personnel.

Eli Bower

Will review and recommend multi-dimensional procedures for assessing child development and child competencies with particular regard for cross-cultural validity and practicality and appropriateness of content. This effort is to be coordinated with that of other panelists to select key dimensions comprising selected constructs and the method for their measurement (e.g., play, games, etc.) All panelists are encouraged to review with Eli Bower the possibility of specifying natural contexts in which more than one dimension (both affective and cognitive) may be observed.

Jean Carew

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) for them in the following area: the social development of Head Start children between the ages of 3 and 7. This work will be coordinated with that of Asa Hilliard. (Construct 2)

Dorothy Eichorn

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) for them in the following area: the gross and fine motor (psychomotor) development of Head Start children, ages 3 to 7. These dimensions should particularly address the high and low ends of the scale, to accommodate those areas most sensitive to program effects. This paper will be coordinated with those of Susan Gordon (Medical) and Jean Sanford (Nutrition). Dorothy Eichorn will arrange for the coordination of these three efforts into a single construct. (Construct 1)

Edward A. De Avila

Will prepare a position paper defining dimensional and measuring elements of cognitive content involved in both construct 3 (cognitive development) and construct 6 (communications competence), from a multi-cultural and developmental aspect. This work will be coordinated with Luis Laosa and Lee C. Lee.

Roger Givens

Will provide input to panelists on the concerns of the Head Start community and the requirements of Head Start children, and also communicate the objectives of the project to Head Start personnel.

William L. Goodwin

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) for them in the following area: motivation among children between 3 and 7 years old, expectations of success and the nature of anticipated reinforcements as they relate generally and specifically to all dimensions of competence. This work will be coordinated with those working on competency constructs (Constructs 6 and 7). William Goodwin will also prepare a paper on "consumer validity for distribution to all panelists".

Susan G. Gordon

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information gathering techniques) for them in the following area: the medical and dental health domain, with reference to Head Start children between the ages of 3 and 7. Particular attention should be given to causative factors that affect overall good health (physical, spiritual, mental), and in particular measures indicating the ability to function under conditions of stress. This paper will be coordinated with those of Dorothy Eichorn and Jean Sanford. Dorothy Eichorn will arrange for the coordination of these three efforts into a single construct. (Construct 1)

Luis Laosa

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) for them in the following area: the perceptual and cognitive development of Head Start children from 3 to 7 years of age. Special reference will be given to the identification of cognitive styles and their implications to this project; specifically, program effects on changing styles (i.e., enhancement of styles), increasing the repertoire of styles available to the child, increases in effectiveness in any particular styles, and a weighting of conditions which affect the styles. Included also will be cognitive content and coping techniques for handling low and high rates of change in sensory inputs. (Construct 3)

Lee C. Lee

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) for them in the following area: the development of communicative competence among children between 3 and 7 years old, together with its impact on development in the social, emotional, cognitive and life-skills domains. Needed adaptations to cross-cultural differences will be defined. (Construct 6)

Winona Sample

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques affecting the functional competencies of children at home, at play and at school; including the range of diverse value systems and life styles of Head Start children. This work will be coordinated with that of Milton Silva in defining specific functional competencies appropriate to all cultures. (Construct 7)

Ruth Marshall

Will provide input to panelists on the concerns of the Head Start community and the requirements of Head Start children, and also communicate the objectives of the project to Head Start personnel.

Jean C. Sanford

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) for them in the following area: the nutritional development of Head Start children from 3 to 7 years of age. The paper will be coordinated with those of Susan Gordon (Medical) and Dorothy Eichorn (psychomotor). Dorothy Eichorn will arrange for the coordination of these three efforts into a single construct. (Construct 1)

Milton Silva

Will prepare a position paper that recommends (1) program-effects dimensions to be measured and (2) appropriate measures (i.e., information-gathering techniques) affecting the functional competencies of children at home, at play and at school. This work will be coordinated with that of Winona Sample in defining specific functional competencies appropriate to all cultures, and with Eli Bower in selecting test situations for gathering multi-dimensional data. (Construct 7)

William Cross

Will prepare a position paper that includes the identification and selections of dimensions of child development and recommend information gathering techniques to measure program effects on the emotional development of Head Start children, ages 3-7. The paper will delineate the difference between self-concept and self-esteem and the application of self-concept development to the social environment and interactions with peers and adults. The paper will include a rationale for these selections and complete supportive documentation. This work will be coordinated with panelists working on affective aspects of other constructs. (Construct 4)

Asa Hilliard

Will review Head Start projects that "appear" to be "working"; these are projects on which there is a general consensus that they are doing "good things" for children. A paper will be prepared defining specifically what these "good things" are, and attempt, if possible, to relate them to measureable dimensions. These outcomes will be communicated to other panelists for consideration in their respective papers; if appropriate, new constructs will be evaluated in accordance with the outline being followed for papers of all other panelists; the objective sensitive to what really works in Head Start.

After this initial phase Asa Hilliard will prepare a paper in conjunction with Jean Carew that recommends (1) program effects to be measured and (2) appropriate measures in the area of social development. (Construct 2)

Note: Ernest Washington will look into the Aesthetic-Moral-Ethnical construct and coordinate with other panelists. (Construct 5)

II. A Discussion of Implications for
Measurement of Program Effects on
Special Needs Children.

EVALUATING THE PROGRESS OF SPECIAL NEEDS CHILDREN

Background and Legal Provision of the Handicapped in Head Start

One of ACYF's long term goals is to have information gathering techniques appropriate to measure program effects on handicapped or special needs children. Developing appropriate measures for use with this sub-population presents usual challenges. The following information provides a perspective for considering this issue as it relates to the project.

The Head Start, Economic Opportunity, and Community Partnership Act of 1974 (P.L. 93-644) required that for fiscal year 1975 not less than ten percent of the total number of enrollment opportunities in Head Start programs in the nation be available for handicapped children, and "That for fiscal year 1976 and thereafter no less than 10 percentum of the total number of enrollment opportunities in each state shall be available for handicapped children..." The Legislation requires that the Secretary shall report to the Congress at least annually on the status of handicapped children being served in Head Start programs, including the numbers of handicapped children being served, their handicapping conditions, and the services being provided.

Definition of the Handicapped

Handicapped children are defined in the Legislation as "mentally retarded, hard of hearing, deaf, speech impaired, visually handicapped, seriously emotionally disturbed, crippled or other health impaired children who by reason thereof require special education and related services." Outside the scope of this definition are children with correctable conditions who do not need special service, or who will not require altered or additional

educational or support services.

Enrollment Figures

Head Start enrollment in Program Year 1976-77 included 36,133 children who were professionally diagnosed as handicapped. This figure represents 13.02% of the total Head Start enrollment. The distribution of handicapped children in Head Start by category of handicap was as follows:

<u>Category</u>	<u>Percent of Handicapped Total</u>
Speech Impaired	50.4
Health Impaired	7.4
Orthopedically Handicapped	7.4
Emotionally Disturbed	6.7
Mentally Retarded	6.6
Specific Learning Disability	5.0
Hearing Impaired	4.4
Visually Impaired	3.9
Deaf	0.4
Blind	0.4

All handicapped children who were enrolled in Head Start programs received the child development services all Head Start children received. One of the most important services which Head Start renders on behalf of the preschool handicapped child is the opportunity to be in a mainstream developmental environment with non-handicapped children. In addition, certain special services assist the family of a handicapped child to maximize the benefits that can be derived from Head Start, and provide a range for a child's special needs through medications and physical facilities or provision of special equipment and materials.

Concern Over Labeling Handicapped Children

In recent years there has been increasing concern expressed about classifying children as handicapped or in need of special services. Despite their fundamental importance, diagnosis and classification pose a dilemma to society. On the one hand, they are required by law in some programs and are needed in research on health and related

social problems and on matters relevant to the planning, administration, delivery, and evaluation of delivery of health and related human services. On the other hand, these processes may result in an individuals acquiring a label that can, and often does, have a serious effect on the child.

Elliot L. Richardson, when Secretary of Health, Education and Welfare, called for a systematic review of the classification and labeling of exceptional children and ten federal agencies joined to sponsor the Project on Classification of Exceptional Children. The final report of this project (The Future of Children, 1975) attempts to convey the significance of classification and labeling; it explores the function of classification in social control; it presents a systematic summary of the state of the art of classifications of handicapped, disadvantaged, and delinquent children and youth; it rejects the currently popular notion that classification can be abandoned and presents a comprehensive plan for linking services with the needs of individual children. It does appear that classification of children with special needs is likely to continue but with increased sensitivity to and efforts at reducing the harmful effects of categories and labels.

It must be remembered that one cannot accept the assumption that all disabled children in any one category function in exactly the same way. Calling a child physically handicapped, visually impaired or any other term provides little information about his functional levels. Such terms do little more than call attention to the fact that he is subject to some deviation from what is considered average or normal. The extent to which this deviation interferes with learning a social behavior or whether the child will respond to a particular intervention strategy is not

reflected in the label. It is the lack of homogeneity among the handicapping conditions that presents the problem for evaluating this subgroup of the Head Start population.

Head Start Classification of Handicapped Children

The Head Start system of classification of handicapped children leans heavily on a medical system which classifies children by pathological descriptions, etiological schemes and, not infrequently, in anatomical terms. Head Start teachers have to make logical inferences from the so-called medical classifications for educational planning. In practice, the child's handicap should not be determined from the label, but from the burden which is imposed on the child when confronted with educational situations that cannot be resolved by reason of body dysfunction or impairment. In the Head Start setting there are at least four educationally relevant factors that must be considered in evaluating the impact of the child's body dysfunction or impairment on functioning in that setting. The impact or burden employed determines the level or degree of handicap. The four factors are :

- a. Motility: The terms intended to encompass the meanings of all modalities of human behavior which facilitate getting about in the environment.
- b. Communication: Is concerned with the organization and utilization of symbolic systems which are the essential ingredients of abstract thinking and problem solving.
- c. Self Concept: Relates to the concept of how the child views himself and how he believes others view him. Poor self concepts that have their origins in body disorder of sufficient complexity and magnitude tend to interfere with adjustment to the school and very often to the general culture.

- d. Social Interaction: Refers to the mutual stimulation of one person by another and the response that follows. It is hardly necessary to make the point that human beings are social creatures. The family, community and school demand specific social behavior.

The severity of the child's handicap will determine the burden it places on the child in social situations and in particular on the above four educationally relevant factors. Children with profound handicapping conditions are likely to enter Head Start programs with greater developmental delays than children with mild handicapping conditions. A profound hearing loss is more debilitating than a mild hearing loss. A child may be labeled as handicapped because of a mild hearing loss but with preferential seating, noise control and appropriate instruction may not be handicapped in the Head Start classroom environment.

Options for Handling the Handicapped in Program Effects Measurement Project

Information gathering techniques being developed for the project should be appropriate for use with handicapped children in the target population. For the purpose of this project, the categories of handicapping conditions used by Head Start for reporting to Congress will be the conditions to consider in clarifying the methodological issues to determine program effects on this subpopulation.

From the background information presented in this paper, it is possible to formulate several options for handling special needs children in this project. One option involves the development of a severity of handicap rating scale. The

burden imposed by the child's body dysfunction or impairment would be rated at the time of pre and post testing. A special loading factor could be assigned to programs with a higher percentage of enrollments of severely handicapped children.

Applied Management Sciences is completing an evaluation of the process of mainstreaming handicapped children into project Head Start and utilized a four point severity rating scale of handicaps. However, there was some indication that "Head Start staff appear to overstate severity of handicaps" (AMS, March, 1978, 5.5).

Another option expands upon the severity of handicapped rating scale and involves the development of a problem proneness of enrolled population in the program. Particular Head Start populations may have more problem children enrolled and the program may make smaller gains but such gains are in reality more significant. This option, however, emphasizes a deficit approach which is likely to be unacceptable to most groups.

Another option for handling the handicapped population in any program would utilize a weighting technique. On the average, 13 percent of the Head Start population is comprised of handicapped children. Programs enrolling more than the average number would be assigned a weighted factor in the determination of program effects. There may be a tendency under this option to label more children as handicapped and accept more handicapped children with less severe conditions.

An analysis of the Head Start handicapped population provides

the rationale that makes the final option discussed more attractive. The option involves treating the handicapped as any other group but providing special instructions for modification of measure in administration.

It should be noted that 50 percent of the handicapped children enrolled in Head Start programs are classified in the speech impaired category. The majority of these cases are diagnosed as articulation disorders that are influenced strongly by maturational factors. The rate of improvement of articulation disorders is rapid during the first two years of school. Very little modification will be required to administer information gathering techniques to this category of handicapped children or those with voice and dysfluency difficulty. Children with delayed speech may require modifications in administration of measures where expressive responses are solicited. This large segment of the handicapped population, the speech impaired, does not present any significant problems for inclusion in the program for measurement of program effects.

The mentally retarded comprise $.858(6.6\% \times .13)$ percent of the Head Start population. Current classification systems of mental retardation depend heavily upon standardized measures of intelligence. However, these tests do not supply information that can be used by a teacher or parent to facilitate determination of program content. Caution should be used in applying the label of mental retardation during the preschool years. It is more useful to consider intelligence as modifiable with changes occurring as a function of the organism - environment interaction. As a result, mental retardation may be considered as reversible.

Rather than emphasizing measurement of static mental functions and knowledge, the modificational approach emphasizes assessment of each individual's learning potential. This concept is compatible with the measurements concepts followed in this project. As a result, the mentally retarded present no unusual problems in being included in evaluation population.

Cerebral palsied children and other low incident categories of the handicapped population, such as the deaf and blind, will require more extensive modification. However, the deaf account for 0.05 percent of the Head Start population and the blind 0.05 percent, (.4% of the handicapped population). When considering the sample included in an overall program evaluation only a very small number of children would require special modification procedures and their impacts on overall program effects would be negligible. For this reason, it is not necessary to select an option for treating this subpopulation in any special manner. Coding procedures could be followed that allow identifying this subgroup for special analysis.

A manual can be prepared that includes directions to para-professionals who will be administering the information gathering techniques. This manual would provide very specific instruction concerning these modifications. High/Scope has developed a Field Procedures Manual for administration of measures used in Project Developmental continuity. This manual includes instructions to test administrators on "situations you may encounter and what to do"; procedures for handling the negative child, the overactive child, the distracted child, and the non-responsive or shy child. Similar directions can be developed for categories

of handicapped children who will be evaluated in the Project Effects Measurement Project.

There are undoubtedly other options available for handling the handicapped in this project. These few options are presented to stimulate discussion on this issue.

III.

APPROACHES TO THE AGGREGATION OF INFORMATION By Richard Wolf

A problem that investigators in all areas of science persistently face is how to aggregate information in ways that yield meaningful indices. Take, for example, the Consumer Price Index. Knowledgeable economists and others faced the problem of determining an index that would indicate how much it was costing people to live. Developing such an index presented a number of formidable problems. What items should be included in the index? How much weight should be given to each item? On what basis should items be added to or deleted from the index? How should the index be expressed? The list could go on and on.

The field of educational testing faces the same problem at a number of levels. How many questions or items of various types should be included in a test? How should test information be aggregated to obtain an index of overall achievement?

The Mediatrix (?) Project also faces the same set of problems. In the case of the Mediatrix Project, the problems are fourfold. First, what are the major constructs that Head Start Programs deal with? Second, what are the major dimensions within each construct? Third, what are suitable measures of the various dimensions? Fourth, how should the information obtained from the use of the measures be combined to determine an overall measure of attainment in the domain that the construct represents? These are all formidable problems. Currently, work is underway that is intended to deal with the first three problems. The purpose of this paper is to deal with the fourth, namely, how should infor-

mation be weighted to obtain an overall index for a construct.

It should be recognized at the outset that the Mediac Project is concerned with program effects. That is, the individual Head Start Programs will be the objects of attention and hence the unit of analysis. While data will be gathered on program participants, these will be aggregated, by measure, for each program. The program results will then need to be aggregated in such a way as to yield indices for the various constructs of interest.

There are two general approaches to the aggregation of information to obtain an overall index of performance. The first is judgmental while the second is empirical. A judgmental approach entails going to people who are deemed knowledgeable and seeking their views as to the weight to be accorded each item (dimension) of information. Two questions that immediately spring to mind about the use of a judgmental approach center on the identification and selection of the people deemed knowledgeable and the ways in which differences in judgements among such knowledgeable people are handled. In the case of the Mediac Project, there are several sets of knowledgeable people that can be readily identified and have already been so identified. A panel of experts in the field of child development has been identified and queried with regard to the objectives of Head Start Programs. Head Start Program Directors and Teachers have been solicited for their views as well as panels of parents of Head Start children. Which is the knowledgeable group? Clearly, a good case could be made for using any one of these groups. However, that isn't the point. The point is that a judgmental approach to aggregating information entails the

use of a knowledgeable group of people and the selection of any single group will have profound consequences on the results obtained. In the case of the MediAx Project, there is already a fair body of evidence to suggest that these various groups differ in their views as to what the objectives of Head Start Programs should be. It would seem safe to presume that they would likely differ on how much weight should be accorded each dimension in deriving an overall index of accomplishment for each domain that Head Start programs address.

This is only part of the story, however. Available information from the various groups noted above indicates that there is considerable variation within each group on what the goals of Head Start Programs should be. It is thus highly likely that there would also be variation within a group as to the weighting of information even if one could decide which knowledgeable group should be given the responsibility of supplying the weights. Reconciling such differences would be a real problem.

The above discussion was intended to expose some of the problems that one faces in executing a judgmental approach to the aggregation of information to obtain an overall index. These problems are painfully real and must be dealt with. The mere existence of problems does not invalidate the approach, however. The judgmental approach has had a long history, some of it rather distinguished. Where there is an agreed upon group of knowledgeable persons and a reasonable degree of agreement among them, then the judgmental approach can work rather well. In the case of the MediAx Project, neither of these two conditions seems to exist.

Hence, it would be questionable to adopt a judgmental approach. Furthermore, not only would the approach probably not work well, but those groups that are not selected to participate in the weighting process might be miffed enough by their rejection to create political problems for the enterprise.

The second major approach to the aggregation of information is empirical. It refers to a variety of procedures that can be used to ascertain a set of weights that are optimal according to various analytic criteria. Different specific empirical procedures will yield different sets of weights depending upon the analytic criteria invoked. It is the responsibility of the investigators to specify the criteria to be used.

While the empirical approach is characterized by the use of statistical procedures, judgment enters into the process at a number of points. The empirical approach is not a mindless one. Judgments need to be made with regard to the major constructs to be studied, the major dimensions that make up the construct and the measures that are to serve as indicators of the dimensions. Furthermore, the particular statistical procedures need to be selected, and criteria for determining what constitutes an optimal set of weights need to be specified. After this is done, the relationships between each measure and the construct needs to be determined so that the statistical model that is specified is the correct one. Judgment must be involved in each of these decisions. Thus, the empirical approach relies heavily on human judgment. The factors that enter into making such judgments are knowledge of the substantive area of the investigation and knowledge of statistical methods.

An empirical approach to the weighting problem can perhaps best be explained by example. The basic model is as follows:

$$S_C = f(A....P) + C$$

where;

S_C = Status on a construct for a particular program, $f(A....P)$ = is a function (f) of dimensions A through P - or whatever number of dimensions that are included for study.

C = error that arises in the study process.

The model is stated in its most general form. Still it is instructive. It states that a program's status on a construct is a function of a number of dimensions plus some error in the measuring process. It is presumed also that the program is the object of attention and that all measures represent group or program performance although they may come from measurements taken on all children enrolled in a program. The model applies equally well to individuals as well as groups, however. In fact, it has typically been used with the individual as the unit of analysis. Of course, there is no reason why between-persons as well as between-groups analysis can't be carried out (see, for example, R. Wolf, Achievement in America. New York: Teachers College Press, 1977).

The weighting problem in the empirical approach involves determining a set of numerical weights to apply to variables A....P that will yield the best estimate of S_C . The term "best" in the last sentence usually means that the set of weights applied to the dimensions results in making C as small as possible, i.e., as close to zero as possible. It is further assumed that the

measurement of dimensions A....P results in a set of numbers along a continuum that indicates more (or less) of a particular dimension. Ideally, the continuum would be an equal interval scale in which the difference between, say, a score of 14 and 15 means the same amount of the dimension as between a score of 21 and 22. While this is the ideal, it is not critical and many worthwhile investigations have been carried out with data that do not meet ideal requirements. One simply has to be a bit more careful. The general requirement that higher scores be associated with more (or less) of a particular dimension, however, holds. It should not be difficult to meet.

The specification of the construct as well as the selection of the dimensions to be included in the analysis is largely a matter of judgment and should result from work currently underway. The resulting dimensions will, however, need to be reviewed before they are included in any empirical investigation. Definitions of the dimensions will require careful scrutiny as well the resulting measures that are intended to get at the dimensions. That work is to be done in the near future. In addition, cost considerations will play a part in the selection of dimensions and measures. Some measures may simply be prohibitive from a cost point of view.

It is to be expected that errors will be made in both the specification of the constructs and the selection of the dimensions that are considered to be pertinent to a construct. That's perfectly normal. There are two such kinds of errors. The first is failing to include a dimension that is germane to a construct. That is

called an error of omission. The second kind of error is including a dimension that is not germane to a construct. That is called an error of commission. Of the two kinds of errors, the first is by far the most serious. If a pertinent dimension is omitted, there will be no way of knowing this from the results of any analysis besides the fact that the error term (e) is sizable. However, a sizable error term will not be an indication of what has been omitted. On the other hand, if a dimension that is not relevant to a construct has been included in an investigation, it will probably yield a zero weight. If so, then it can be disregarded. Thus, a general strategy would be to include as many dimensions as possible or, practically speaking, all those dimensions that anyone associated with the project thinks bears a relationship to a construct.

The next major point involves the construct measures (S_c). Either one has a measure of S_c or one does not. If one has a measure of S_c , one kind of analytic procedure can be used (multiple regression analysis or multiple discriminant analysis, depending on the nature of the measure of S_c) while if there is no measure of S_c , another statistical procedure (usually factor analysis) must be used. It is clearly preferable to have a measure of S_c .

Obtaining an adequate measure of S_c for a program will be a difficult undertaking. However, the time and effort would seem well worth it. Ideally, assessment teams of specially trained people would be needed to visit a sample of Head Start Programs and make a determination of S_c on the spot. This is what accrediting bodies typically do. While the task is not easy, it can be done and is done routinely with high schools, colleges and graduate

professional programs.

The selection of a sample of programs for inclusion in a study is a straightforward if not an easy task. The goal of sampling would be to select a representative group of programs from the population of Head Start programs. A sample of 300 programs would be most satisfactory. For study purposes, it would be wise, regardless of whatever analytic procedure is used, to split the total sample of 300 programs into two equivalent groups. These can be termed Sample A and Sample B although they are, in effect, subsamples. The use of the two sample will be discussed later.

At this point, assume that a construct and a set of dimensions have been identified. Further, two equivalent samples of programs have been selected. The data collection phase of the enterprise entails visits to all selected programs and obtaining measures of all dimensions A....P as well as S_c for each program. The resulting data matrix can be represented as follows:

Program #	<u>Dimensions</u>		S_c
	A.....P		
1	X_{A1}	 X_{P1}	S_{c1}
2	X_{A2}	 X_{P2}	S_{c2}
3			
N	X_{AN}	X_{PN}	S_{cN}

The analytic problem is to now determine a set of weights to apply to the measures and a way of combining these products (weight times measure) to obtain the best estimate of S_c for all programs in a sample. The usual way of doing this is through multiple

regression analysis. The most common form of multiple regression analysis is an additive model that is:

$$S_{ci} = b_1X_{A1} + \dots + X_{p1}$$

While this is the most common form of the regression model, it is not necessarily correct. It's chief claim to widespread use is its simplicity and parsimony. Its simplicity stems from the fact that it assumes an additive relationship of the dimension while its parsimony inheres in the fact that the relationship of dimensions with the construct is linear. These, of course, are appealing features of a model. That, however, doesn't make them correct.

A conservative approach would be to first study the relationship of each dimension with the construct to ascertain its nature--linear or non-linear. Depending on the outcome, appropriate transformations can be made, e.g., adding square, cubic, quadratic terms, etc. After this is done, interaction effects among variables can be studied and, if necessary, product terms can be added. The point is, that in the face of uncertainty, considerable exploratory work needs to be carried out before a model can be properly specified. Specification of a linear additive model has the appeal of simplicity but little else to recommend it. It would seem that given the complexity of the constructs and dimensions being studied a considerable amount of exploratory work will be needed to establish an appropriate model.

Such exploratory work and the eventual testing of the model using multiple regression analysis (multiple discriminant or canonical analysis if S_c cannot be measured on a single continuous

scale) would be carried out on one sample only. The other sample would not be touched. Only after an adequate model has been determined would the other sample be used. The results of the first analysis would then be applied to the second sample to determine the stability of results. Better yet, independent work on both studies can be done and the results from each sample analysis could then be applied to the other sample after comparison. This technique is called double cross-validation and was originally developed by Mosier (Educational and Psychological Measurement, Summer, 1971). Whatever approach to validation is adopted, it will be necessary to check out the results to determine if they represent genuine information about the weights to apply to the dimensions and the proper way to weigh the information or whether they are primarily artifacts arising out of a particular body of data.

The above discussion has been necessarily terse. The main purpose of this paper has been to present and discuss approaches to the aggregation of information for the Medias Project. It has not been the purpose of this paper to present a detailed discussion of a body of analytic procedures. This can be found in any number of standard works. Thus, this paper can be viewed as a general guide instead of a detailed road map.

There are two general approaches to the weighting of information. One is judgmental while the other is empirical. Neither is a pure approach. Some empirical procedures can be used as a judgmental approach, largely in connection with the pooling of judgments. However, it is this writer's opinion that the Medias Project will not be well served by such an approach for the reasons cited earlier. Simply restated, no one knows enough to pull it

off. Accordingly, it has been recommended that an empirical approach be followed. This approach requires careful work by the investigators. Considerable judgment is required in the identification of the constructs to be studied, the selection of dimensions to be included and the appraisal of specific measures of dimensions. Furthermore, the specification of the construct and the assessment by a special team needs to be done judiciously. Special expertise will also be needed to determine the most suitable model for the analysis. This will involve a study of binariate relationships of each dimension with the construct as well as exploratory work regarding the best way to aggregate the variables, including the introduction of interaction terms into the model. Thus, while the empirical approach makes heavy use of statistical procedures, it must do so in a judicious way.